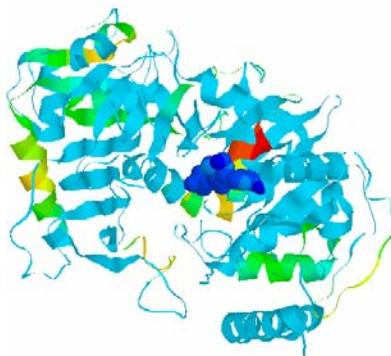


# **Identification of ATP – Binding Proteins within Sequenced Bacterial Genomes Using Random Peptide Phage Display**

**Lee Makowski, Suneeta Potluri, Fred J. Stevens  
and Diane J. Rodi**

**Combinatorial Biology Unit,  
Biosciences Division,**



## **Consensus Sequences as clues to protein function**

qgtrtiqalnnvslhvpagqiyyvgivgasgagkstlircvnllerptegsvlvdggelttl  
Ksfgpvhalksvnltvypgeihallgengagkstlmkvlsqiheptkgtitinnisynkl  
Ktfpgvkaltdisfdcyagqvhalmgengagkstllkilsgnyapttgsvvinqemsfs  
hqalstqlldaimpycgntlrlgvgtgtpgagkstfleafgmllicreglkvaviavdpssp  
aekrniflvpgpmgagkstigrqlaqqlnmefydsdqeiakrtga  
cfygahqalfditldcpqgetlvlrgpsgagkssllrvlnllemprsgtlniaghfdft  
qdvaestrlgplsgevrageeilhlvgpngagkstllarmagmtsgkgsiqfaggqpleaws

## P-loop Structures

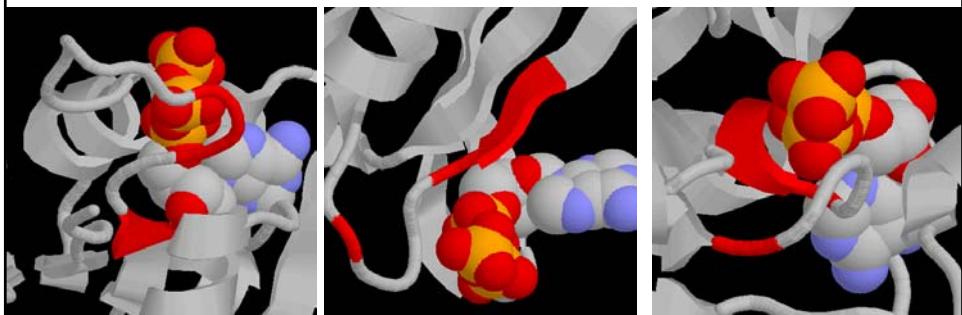
*The most common small-molecule binding motif...*

1mjh hypothetical protein

1nsf n-ethylmaleimide sensitive

2gnk nitrogen regulatory protein

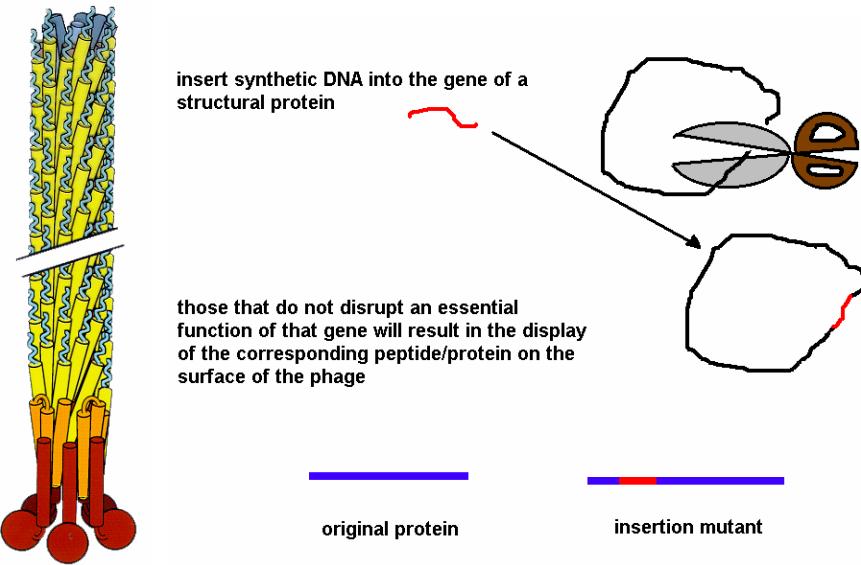
**G-xxxx-GK[ST]**



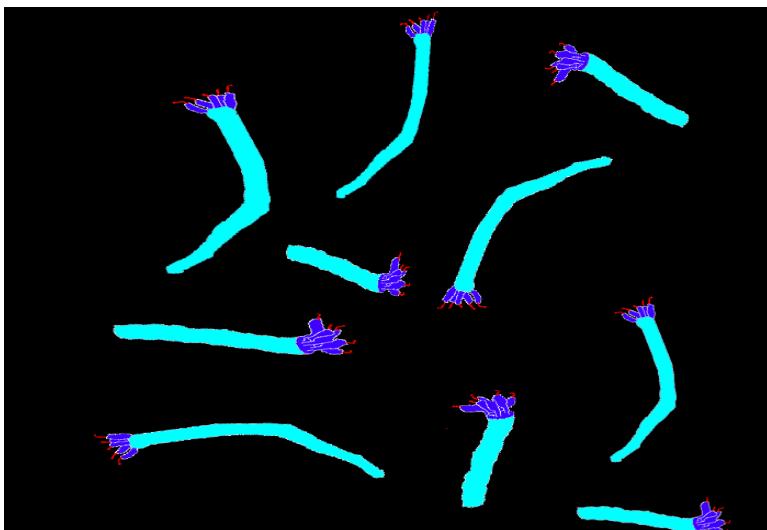
**Can we find a strategy for identifying new small molecule consensus sequences – even for those small molecules that bind to only a relatively small number of proteins?**

**For instance, could we produce a lot of different proteins with affinity to the small molecule and then characterize the binding...**

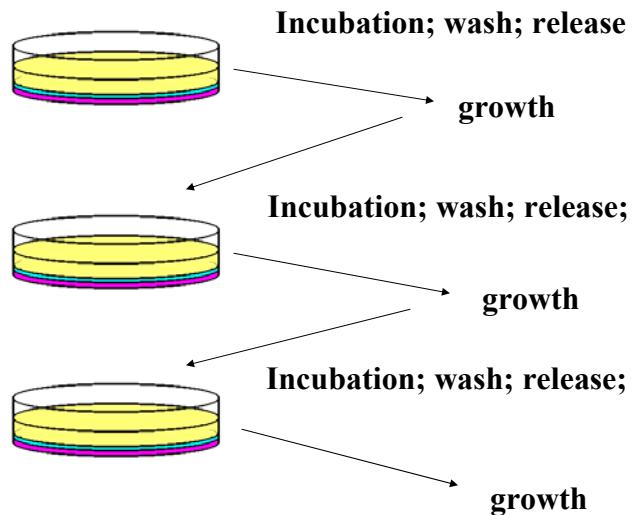
# Phage Display Technology



$20^{12}$  ( $4 \times 10^{15}$ ) possible phage particles  
- each one different



## Biopanning

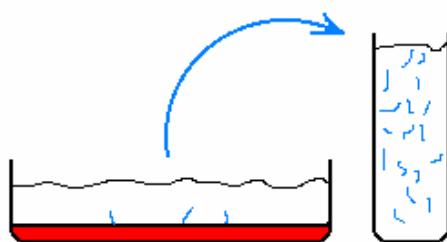


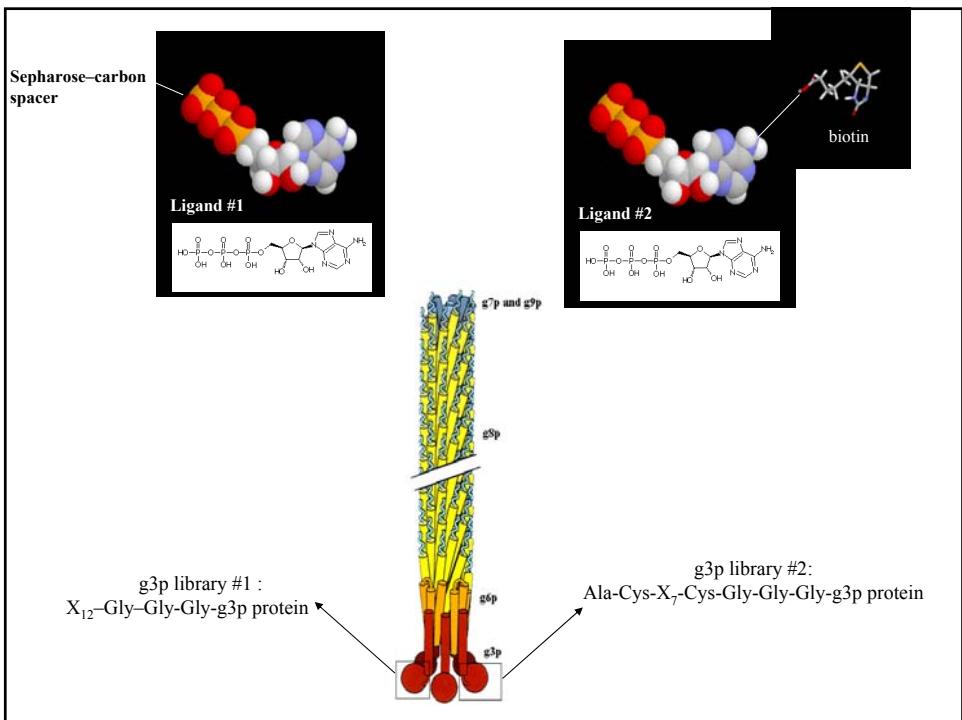
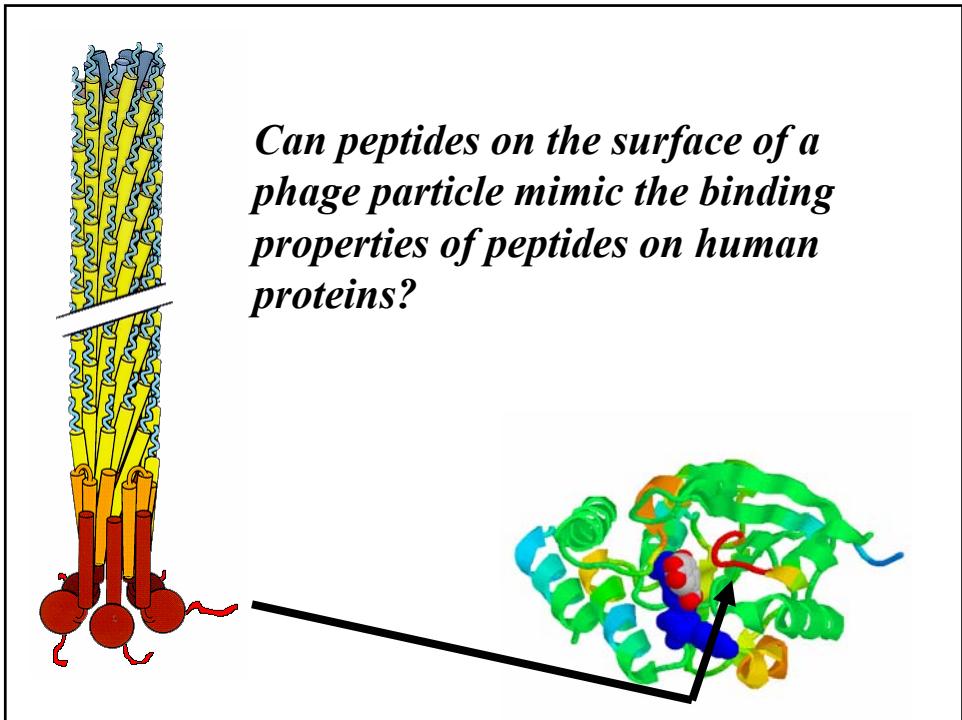
## Phage Display:

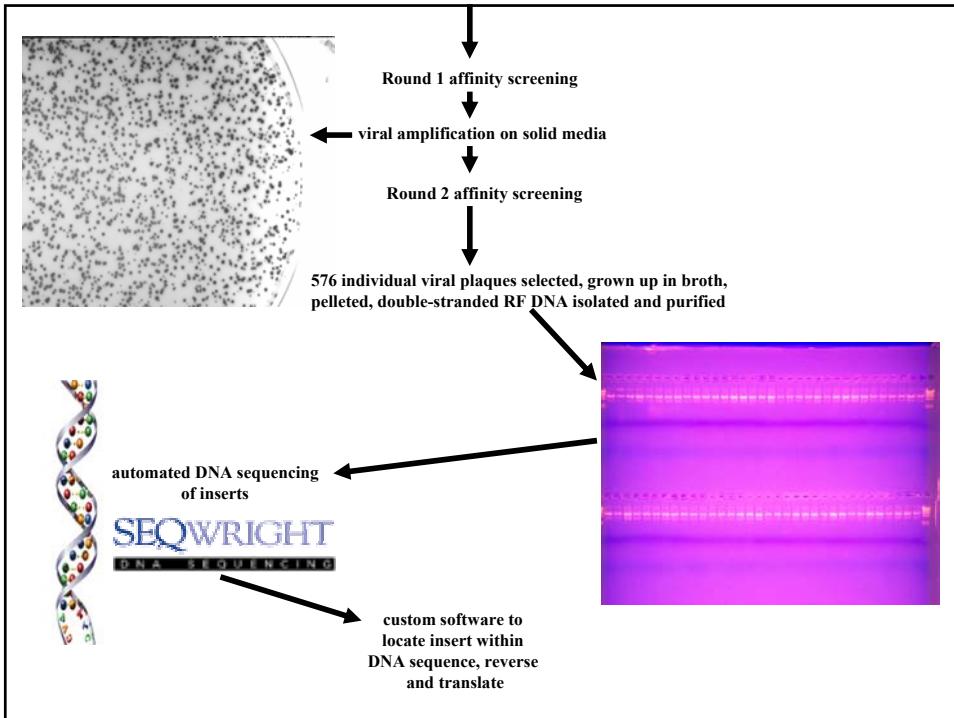
**advantage over traditional affinity methods:**

**isolate one binding phage -**

**grow up a barrel to characterize**

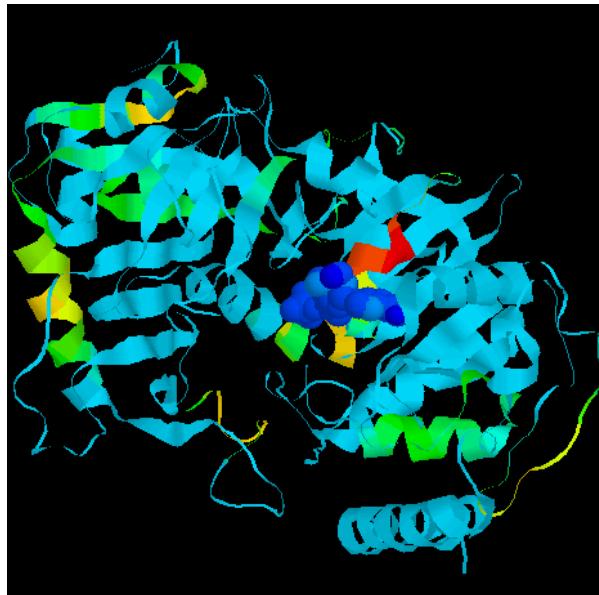






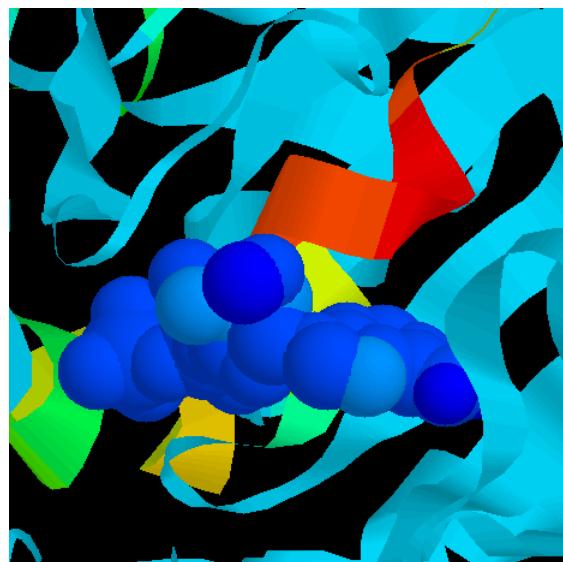
## **$\gamma$ – ATP-Sepharose-binding peptides**

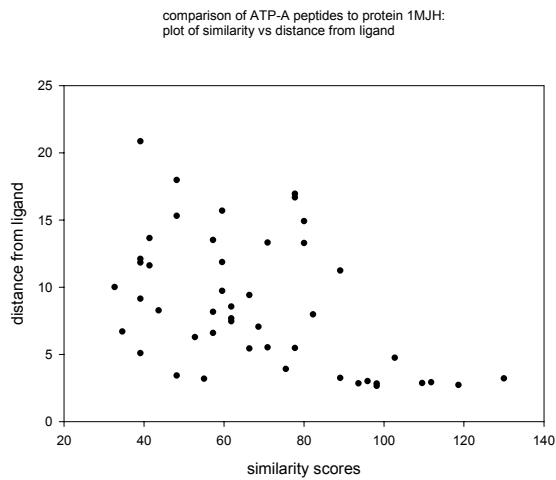
### Phosphoenolpyruvate carboxykinase



Red=high similarity

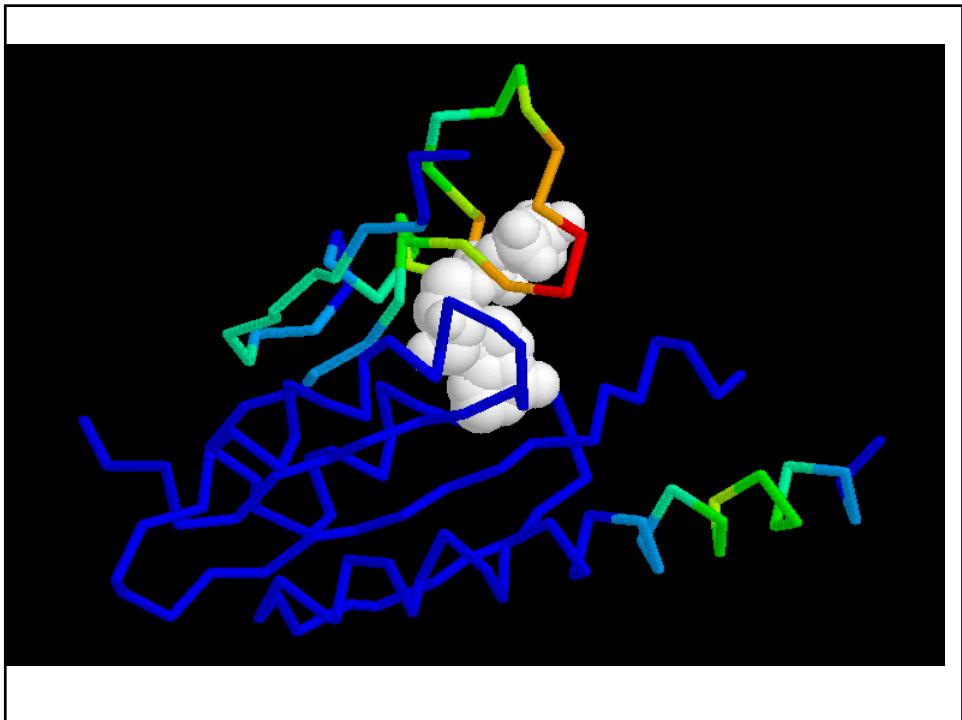
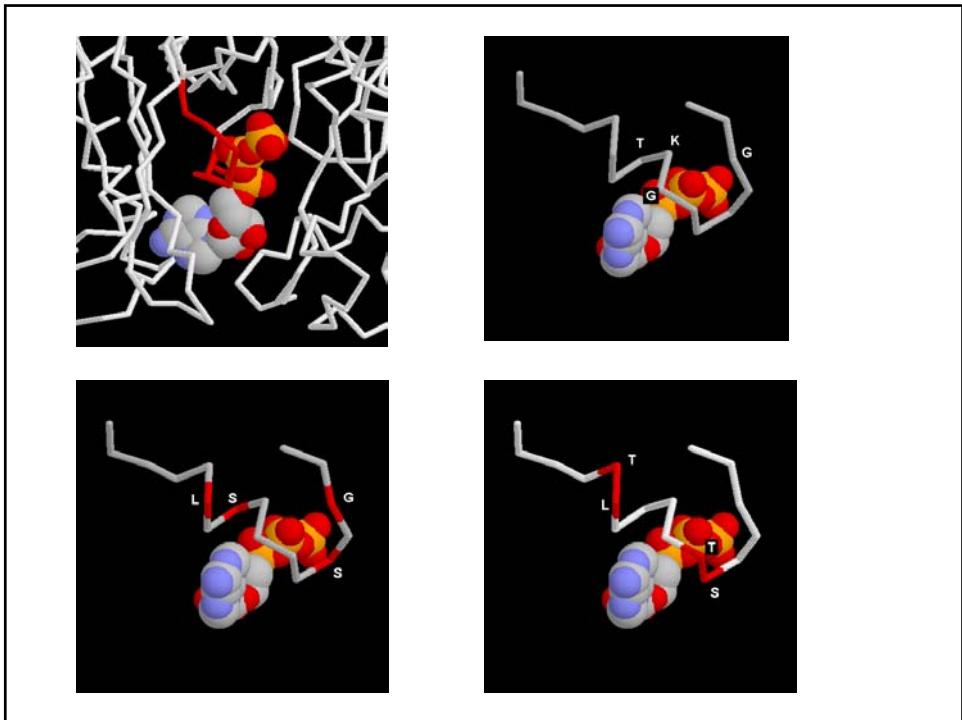
Blue=low similarity





xxxxxxxxxxxxxx xxx  
k<sub>240</sub>gdvavffglsgtgkttlstdpkrrligddehgwddgvgfnfeggc<sub>285</sub>  
tfvtssstdtrrs  
idmqktnlahgp  
qttlnsdfprrtr  
lrvpppllsvnpr  
sssfittlsgpr

## Alignment of sub-population of ATP-binding peptides against 1AYL sequence [XXX = ATP-contact residues]



- 1 22.17>gi|1787942|gb|AAC74725.1| member of ATP-dependent helicase superfamily II  
     [Escherichia coli K12]  
 2 21.73>gi|2367341|gb|AAC77025.1| diadenosine tetraphosphatase [Escherichia coli K12]  
 3 21.43>gi|1787197|gb|AAC74049.1| methylglyoxal synthase [Escherichia coli K12]  
 4 21.13>gi|1787702|gb|AAC74514.1| putative virulence protein [Escherichia coli K12]  
 5 21.13>gi|1786345|gb|AAC73262.1| ATP-binding component of hydroxymate-dependent iron transport  
 6 20.98>gi|1787668|gb|AAC74483.1| split orf [Escherichia coli K12]  
 7 20.68>gi|1789552|gb|AAC76195.1| tryptophan-specific transport protein [Escherichia coli K12]  
 8 20.24>gi|1787121|gb|AAC73980.1| anaerobic dimethyl sulfoxide reductase subunit A  
 9 20.24>gi|1788092|gb|AAC74861.1| putative amino acid/amine transport protein  
 10 20.09>gi|1787885|gb|AAC74673.1| putative transport protein [Escherichia coli K12]  
 11 20.09>gi|1788863|gb|AAC75568.1| orf, hypothetical protein [Escherichia coli K12]  
 12 19.35>gi|1788112|gb|AAC74880.1| orf, hypothetical protein [Escherichia coli K12]  
 13 19.35>gi|1787051|gb|AAC73916.1| putative ATP-binding component of a transport system  
     [Escherichia coli K12]  
 14 19.20>gi|1790393|gb|AAC76938.1| phosphoenolpyruvate carboxylase [Escherichia coli K12]  
 15 18.90>gi|1789971|gb|AAC76573.1| 3-methyl-adenine DNA glycosylase I, constitutive  
     [Escherichia coli]  
 16 18.90>gi|1789962|gb|AAC76565.1| putative ATP-binding component of dipeptide transport system  
 17 18.90>gi|1789453|gb|AAC76107.1| aerotaxis sensor receptor, flavoprotein [Escherichia coli K12]  
 18 18.90>gi|1789230|gb|AAC75904.1| orf, hypothetical protein [Escherichia coli K12]  
 19 18.75>gi|1788472|gb|AAC75210.1| ATP-binding component of methyl-galactoside transport  
 20 18.75>gi|1786804|gb|AAC73690.1| ferric enterobactin transport protein [Escherichia coli K12]  
 21 18.30>gi|1790736|gb|AAC77240.1| IS30 transposase [Escherichia coli K12]  
 22 18.30>gi|1788909|gb|AAC75610.1| phosphoribosylformyl-glycineamide synthetase =  
     FGAM synthetase  
 23 18.15>gi|1788344|gb|AAC75093.1| putative glucose transferase [Escherichia coli K12]  
 24 18.01>gi|1789286|gb|AAC75956.1| putative enzyme [Escherichia coli K12]  
 25 18.01>gi|1788151|gb|AAC74916.1| orf, hypothetical protein [Escherichia coli K12]

- 26 17.86>gi|1790375|gb|AAC76921.1| cystathione gamma-synthase [Escherichia coli K12]  
 27 17.86>gi|1787318|gb|AAC74162.1| flagellar biosynthesis, cell-distal portion of basal-body rod [ Escherichia coli K12]  
 28 17.86>gi|1790611|gb|AAC77126.1| N-acetylmuramoyl-l-alanine amidase II; a murein hydrolase [Escherichia coli K12]  
 29 17.86>gi|1789350|gb|AAC76014.1| glycolate oxidase iron-sulfur subunit [Escherichia coli K12]  
 30 17.71>gi|1786336|gb|AAC73254.1| poly(A) polymerase I [Escherichia coli K12]  
 31 17.56>gi|1790857|gb|AAC77349.1| right origin-binding protein [Escherichia coli K12]  
 32 17.56>gi|1786801|gb|AAC73687.1| ATP-dependent serine activating enzyme  
 33 17.41>gi|1790472|gb|AAC77009.1| chorismate lyase [Escherichia coli K12]  
 34 17.41>gi|1786521|gb|AAC73430.1| orf, hypothetical protein [Escherichia coli K12]  
 35 17.26>gi|2367309|gb|AAC76838.1| orf, hypothetical protein [Escherichia coli K12]  
 36 17.26>gi|1788767|gb|AAC75480.1| orf, hypothetical protein [Escherichia coli K12]  
 37 17.26>gi|1788524|gb|AAC75256.1| cytochrome c-type biogenesis protein [Escherichia coli K12]  
 38 17.26>gi|1787547|gb|AAC74372.1| putative ATP-binding protein of peptide transport system  
     [Escherichia coli K12]  
 39 17.26>gi|1786852|gb|AAC73734.1| a minor lipoprotein [Escherichia coli K12]  
 40 17.26>gi|1786549|gb|AAC73456.1| putative transport protein [Escherichia coli K12]  
 41 17.11>gi|1789345|gb|AAC76009.1| orf, hypothetical protein [Escherichia coli K12]  
 42 17.11>gi|1787529|gb|AAC74356.1| DNA topoisomerase type I, omega protein [Escherichia coli K12]  
 43 16.96>gi|1789336|gb|AAC76001.1| transport of nucleosides, permease protein [Escherichia coli K12]  
 44 16.96>gi|1787906|gb|AAC74692.1| repressor of malX and Y genes [Escherichia coli K12]  
 45 16.82>gi|2367260|gb|AAC76690.1| regulator of uhpT [Escherichia coli K12]  
 46 16.67>gi|1790745|gb|AAC77248.1| regulator for fec operon, periplasmic [Escherichia coli K12]  
 47 16.67>gi|1788557|gb|AAC75286.1| orf, hypothetical protein [Escherichia coli K12]  
 48 16.67>gi|1787909|gb|AAC74694.1| enzyme that may degrade or block biosynthesis of endogenous  
     maltozose [Escherichia coli K12]  
 49 16.67>gi|1787771|gb|AAC74568.1| orf, hypothetical protein [Escherichia coli K12]  
 50 16.52>gi|1790835|gb|AAC77328.1| peptide chain release factor RF-3 [Escherichia coli K12]

## Conclusions

- (i) *Similarity to sequences of affinity-selected peptides is predictive*
- (ii) *Not all motifs represented in single library*
- (iii) *Not all motifs obtained with single target*
- (iv) *Specificity of motifs may depend on details of target used (e.g. ATP/GTP)*
- (v) *Possible to predict probability of any given proteins binding of target (?)*

## P-loop Structures

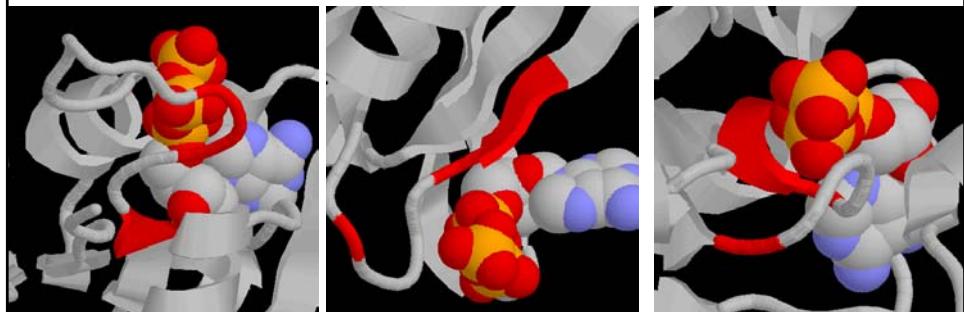
*The most common small-molecule binding motif...*

1mjh hypothetical protein

1nsf n-ethylmaleimide sensitive

2gnk nitrogen regulatory protein

**G-xxxx-GK[ST]**



**Ligand Binding Sites are Composed of Stable Foundations overlayed by Structurally Unstable, Flexible Loops**

